Statistiques et Probabilité

I. Statistiques :

→ Les statistiques sont des méthodes pour collecter, présenter, et analyser des données.

Exemple : Etude de l'âge des étudiants d'une université.

prenons par exemple 10 étudiants

$P = \{19, 21, 25, 19, 22, 21, 25, 19, 20, 20\}$.

x_i	19	20	21	22	25
n_i	3	2	2	1	2
$n_i\uparrow$	3	5	7	8	10
f_i	30%	20%	20%	10%	20%
$f_i\uparrow$	30%	50%	70%	80%	100%

- L'âge : variable stitistique OU Caractère à étudier.
- Les Etudiants : Population, chacun d'eux s'appelle un individu.
- Card (P) = $N = \sum_{i=1}^{n} N_i$: L'effective totale
- modalité : la valeur prise pour notre car. stitistique. (x_i).

$n_i\uparrow$: Les effectifs croissantes.

$f_i\uparrow$: Les fréquences croissantes.

$$f_i = \frac{n_i}{N}$$

- Types des variables statistiques

1/ Caractères qualitatifs :

ce sont des caractères dont les modalités s'échappent de la mesure, les peuvent seulement être constatés.

2/ Caractères quantitatifs :

→ discrètes : Si x prend des valeurs naturelles.

→ Continus : Si x prend des valeurs réelles.

- La moyenne (L'espérence Math.)

$$E(X) = \overline{X} = \frac{1}{N} \cdot \sum_{i=1}^{n} n_i x_i$$

- La Variance :

$$V(X) = E(X^2) - E(X)^2$$

$$V(X) = \frac{1}{N} \cdot \sum n_i (x_i - \overline{X})^2$$

- L'ecart-type :

$$\sigma(X) = \sqrt{V(X)}$$

Scanné avec CamScanner

Remarque : L'écart type nous permet de savoir si les valeurs de la modalité sont proches ou non.

$\sigma(X) = 0 \Rightarrow$ les valeurs sont proches.

Sinon $\Rightarrow$ les valeurs sont mal-distribuées.

- **Types des moyennes :**

→ moyenne harmonique :

$$\frac{1}{H} = \frac{1}{n} \sum_{i=1}^{k} \frac{n_i}{x_i}$$

→ moyenne géométrique :

$$G = \sqrt[n]{\prod_{i=1}^{k} x_i^{n_i}}$$

ou :

$$G = \prod_{i=1}^{k} x_i^{f_i}$$

ou :

$$\ln(G) = \sum_{i=1}^{k} f_i \cdot \ln(x_i)$$

→ moyenne quadratique :

$$Q = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} n_i (x_i)^2}$$

Remarque : quelle-que soit série statistique considérée on a : $H \leqslant G \leqslant \bar{X} \leqslant Q$

- **La médiane :**

→ cas d'une var. discrète :

⚠ on doit d'abord ordonner la série

impair	pair
$N = 2p+1$	$N = 2p$
$m_e = x_{p+1}$	$m_e = \frac{x_p + x_{p+1}}{2}$

→ **Cas d'une var. continue :**

Dans ce cas les effectifs n'ont pas besoin d'être ordonnés.

- Si les amplétudes sont égaux : la classe médiane est la classe $[e_{i-1}, e_i[$ dont la fréquence cumulée croissante approche 0,5 par excès càd $F_i \geq 0,5$.

$$m_e = e_{i+1} + (e_i - e_{i-1}) \left(\frac{0,5 - F_{i-1}}{F_i - F_{i-1}} \right)$$

$$= e_{i+1} + (e_i - e_{i-1}) \left(\frac{0,5 - F_{i-1}}{f_i} \right)$$

Scanné avec CamScanner

Le mode :

→ Cas d'une variable discrète

Le mode est la valeur de la modalité qui correspond à l'effectif le plus grand.

- Remarque :
- lorsque la distribution possède :

un seul mode → unimodale

2 modes → bimodale.

2 modes consécutifs → intervalle modal.

→ Cas d'une variable Continue

$$m_o = e_{io-1} + (e_{io} - e_{io-1})\left(\frac{d_1}{d_1+d_2}\right)$$

avec e_{io} et e_{io-1} sont les extremités de la classe modale.

et $d_1 = f_i - f_{i-1}$ $d_2 = f_i - f_{i+1}$

lorsque les classes ont la même amplitude et $d_1 = f_i' - f_{i-1}'$

et $d_2 = f_i' - f_{i+1}'$ en cas des amplitudes différentes.

avec $f_i' = f_i \times \frac{a_r}{a_i}$ s'appellent les fréquences corrigées.

a_r : l'amplitude la plus petite.

a_i : l'amplitude de la classe.

- les Quartiles :

→ Cas d'une variable Continue :

- même chose pour la médiane on change juste le 0,5 par 0,25 pour le Q_1 et 0,75 pour le Q_3. (avec Q_2 = médiane).

$$Q_1 = e_{i+1} + (e_i - e_{i-1})\left(\frac{0,25 - F_{i-1}}{F_i - F_{i-1}}\right)$$

$$Q_3 = e_{i+1} + (e_i - e_{i-1})\left(\frac{0,75 - F_{i-1}}{F_i - F_{i-1}}\right)$$

(avec $f_i = F_i - F_{i-1}$)

⚠ et En cas des amplitudes inégaux, on change les fréquences relatives par les fréquences corrigées.

Notions Principales :

- Coeff de variance : $C_v = \frac{\sigma(x)}{E(x)}$
- L'écart Moyen : $\frac{1}{n}\sum_{i=1}^{n} |x_i - \bar{x}|$
- L'écart Absolu Moyen :

→ Var. Discrète : $\frac{1}{N}\sum n_i |x_i - \bar{x}|$

→ Var. Continue : $\frac{1}{N}\sum_{i=1}^{n} n_i |c_i - \bar{x}|$

avec c_i sont les centres des intervalles $[e_{i-1}, e_i[= \frac{e_{i-1}+e_i}{2}$

Scanné avec CamScanner

II. Probabilités :

1. Combinaison de k objets

• Si on a n objets, toute suite composée de k objets parmi n s'appelle une combinaison de k objets.

$$C_n^k = \frac{n!}{k!\,(n-k)!} = C_n^{n-k}$$

$$= C_{n-1}^{k} + C_{n-1}^{k-1}$$

$$C_n^k = C_n^{n-k} \quad ; \quad C_n^k = \frac{A_n^k}{k!}$$

$$C_n^0 = 1 \quad ; \quad C_n^1 = n \quad ; \quad C_n^{n-1} = n$$

2. Arrangement de k objets :

on a n objets, toute suite ordonnée de k parmi n objets s'appellent arrangement de k objets et on a :

$$A_n^k = \frac{n!}{(n-k)!} = A_{n-1}^k + k \cdot A_{n-1}^{k-1}$$

$$A_n^0 = 1 \quad ; \quad A_n^1 = n \quad ; \quad A_n^{n-1} = n!$$

3. Permutation de n objets :

Si on considère n objets, le nombre de permutations possibles est $P_n = n!$

4. Intersection de A et B :

on considère deux événements A et B. l'intersection des deux est notée $(A \cap B)$.

Si $A \cap B = 0 \Rightarrow$ A et B sont incompatibles

5. Réunion de A et B :

c'est l'evenement noté $(A \cup B)$ tel que :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

• **Probabilité conditionnelle :**

la probabilité de A sachant que B a été réalisé est :

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

• $P(\bar{A}/B) = 1 - P(A/B)$.

• $P(A \cup B/C) = P(A/C) + P(B/C)$

→ Si $A \cap B \neq 0$.

• $P(A \cup B/C) = P(A/C) - P(A \cap B/C)$

→ Si $A \cap B \neq 0$.

Scanné avec CamScanner

L'indépendance de A et B :

- A indépendant de B

$$\Rightarrow \begin{cases} P(A/B) = P(A) \\ P(A \cap B) = P(A) \times P(B). \end{cases}$$

$\Rightarrow$ A et $\overline{B}$ sont indépendants.

$\Rightarrow$ $\overline{A}$ et B " "

$\Rightarrow$ $\overline{A}$ et $\overline{B}$ " ".

- **Variable Aléatoire :**

Exp : Le jet d'un dé peut prendre 6 valeurs.

X	1	2	3	4	5	6
P(x)	1/6	1/6	1/6	1/6	1/6	1/6

- L'expérience mathématique :

$$E(x) = \sum x \cdot P(x)$$

- La variance :

$$V(x) = E(x^2) - E^2(x)$$

- L'écart-type :

$$\sigma(x) = \sqrt{V(x)}$$

- Fonction de répartition :

$$F(x) = P(X \leq x) = \sum^{x} P(x)$$

X	1	2	3	4	5	6
P(x)	1/6	1/6	1/6	1/6	1/6	1/6
F(x)	1/6	2/6	3/6	4/6	5/6	1

- **Loi de Bernoulli** : c'est une expérience aléatoire qui a 2 résultats possibles.

  - Succès : p
  - Echec : $q = 1 - p$.

X	0	1
P(x)	q	p.

- **Loi Binomiale :**

- Elle est souvent utilisé dans les cas de plusieurs expérience identiques et indépendantes.

  - (p) Succès
  - (q) Echec.

- Le X caractérise le nb de succès au cours de n expériences aléatoires. $(0 \leq X \leq n)$

$$P(X = x) = C_n^x \cdot p^x \cdot q^{n-x}$$

$$E(x) = np \;,\; V(x) = npq \;,\; \sigma(x) = \sqrt{V(x)}$$

Scanné avec CamScanner